PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/57, 9/52, A61K 38/48, C07K 16/40, C12N 1/21, 5/10	A1	(11) International Publication Number: WO 99/15675 (43) International Publication Date: 1 April 1999 (01.04.99)

(21) International Application Number: PCT/US98/20186

(22) International Filing Date: 24 September 1998 (24.09.98)

(30) Priority Data:
60/059,907 24 September 1997 (24.09.97) US

(71) Applicants *(for all designated States except US)*: REGENTS OF THE UNIVERSITY OF MINNESOTA [US/US]; 100 Church Street Southeast, Minneapolis, MN 55455 (US). AMERICAN CYANAMID COMPANY [US/US]; 5 Giralda Farms, Madison, NJ 07940 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: HOSTETTER, Margaret, K. [US/US]; Ezra Stiles College, 19 Tower Parkway, New Haven, CT 06511 (US). FINKEL, David, J. [US/US]; Apartment 303, 319 8th Street Southeast, Minneapolis, MN 55414 (US). CHENG, Qi [CN/US]; 11240 48th Avenue North, Plymouth, MN 55442 (US). GREEN, Bruce, A. [US/US]; 49 Northfield Gate, Pittsford, NY 14534 (US). MASI, Amy, W. [US/US]; 326 Grand Circle, Caledonia, NY 14423 (US).

(74) Agent: MUETING, Ann, M.; Mueting, Raasch & Gebhardt, P.A., P.O. Box 581415, Minneapolis, MN 55458-1415 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: HUMAN COMPLEMENT C3-DEGRADING PROTEINASE FROM *STREPTOCOCCUS PNEUMONIAE*

```
  1 ATGTCAAGCC TTTTACGTGA ATTGTATGCT AAACCCTTAT CAGAACGCCA
 51 TGTAGAATCT GATGGTCTTA TTTTCGACCC AGCGCAAATC ACAAGTCGAA
101 CCGCCAATGG TGTTGCTGTA CCGCACGGAG ACCATTATCA CTTTATTCCT
151 TATTCACAAC TGTCACCTTT GGAAGAAAAA TTGGTCGTAT TATTCCCCTT
201 CGTTATCGTT CAAACCATTG GGTACCAGAT TCAAAGACCA GAACAACCAG
251 TCCACAATCG ACTCCGGGAA CCTAGTCCAA GTCCGAAACC TGCACCAAAT
301 CCTCAACCAG CTCCAAGCAA TCCAATTGAT GAGAAATTGG TCAAAGAAGC
351 TGTTCGAAAA GTAGGCGATG GTTATGTCTT TGAGGAGAAT GGAGTTGCCT
401 CGTTATATCC CAAGCCAAGG ATCTTACAGC AGAAACAGCA GCAGGCATTG
451 ATAGCAAACT GGCCAAGCAG GAAAGTTTAT CTCATAAGCT AG
```

(57) Abstract

The present invention relates to the identification and use of a family of human complement C3-degrading proteinases expressed by *S. pneumoniae*. The proteinase has a molecular weight of about 15 kD to about 25 kD. A preferred proteinase of this invention includes the amino acid sequence of SEQ ID NO:2.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

HUMAN COMPLEMENT C3-DEGRADING PROTEINASE FROM *STREPTOCOCCUS PNEUMONIAE*

Field of the Invention

This invention relates to *Streptococcus pneumoniae* and in particular this invention relates to the identification of an *S. pneumoniae* protein that is capable of degrading human complement protein, C3.

Background of the Invention

Respiratory infection with the bacterium *Streptococcus pneumoniae* (*S. pneumoniae*) leads to an estimated 500,000 cases of pneumonia and 47,000 deaths annually. Those persons at highest risk of bacteremic pneumococcal infection are infants under two years of age, individuals with a compromised immune system and the elderly. In these populations, *S. pneumoniae* is the leading cause of bacterial pneumonia and meningitis. Moreover, *S. pneumoniae* is the major bacterial cause of ear infections in children of all ages. Both children and the elderly share defects in the synthesis of protective antibodies to pneumococcal capsular polysaccharide after either bacterial colonization, local or systemic infection, or vaccination with purified polysaccharides. *S. pneumoniae* is the leading cause of invasive bacterial respiratory disease in both adults and children with HIV infection and produces hematogenous infection in these patients (Connor et al. *Current Topics in AIDS* 1987;1:185-209 and Janoff et al. *Ann. Intern. Med.* 1992;117(4):314-324).

Individuals who demonstrate the greatest risk for severe infection are not able to make antibodies to the current capsular polysaccharide vaccines. As a result, there are now four conjugate vaccines in clinical trial. Conjugate vaccines consist of pneumococcal capsular polysaccharides coupled to protein carriers or adjuvants in an attempt to boost the antibody response. However, there are other potential problems with conjugate vaccines currently in clinical trials. For example, pneumococcal serotypes that are most prevalent in the United States are different from the serotypes that are most common in places

such as Israel, Western Europe, South Africa, or Scandinavia. Therefore, vaccines that may be useful in one geographic locale may not be useful in another. The potential need to modify currently available capsular polysaccharide vaccines or to develop protein conjugates for capsular vaccines to suit geographic serotype variability entails prohibitive financial and technical complications. Thus, the search for immunogenic, surface-exposed proteins that are conserved worldwide among a variety of virulent serotypes is of prime importance to the prevention of pneumococcal infection and to the formulation of broadly protective pneumococcal vaccines. Moreover, the emergence of penicillin and cephalosporin-resistant pneumococci on a worldwide basis makes the need for effective vaccines even more exigent (Baquero et al. *J. Antimicrob. Chemother.* 1991;28S;31-8).

Several pneumococcal proteins have been proposed for conjugation to pneumococcal capsular polysaccharide or as single immunogens to stimulate immunity against *S. pneumoniae*. Surface proteins that are reported to be involved in adhesion of *S. pneumoniae* to epithelial cells of the respiratory tract include PsaA, PspC/CBP112, and IgA1 proteinase (Sampson et al. *Infect. Immun.* 1994;62:319-324, Sheffield et al. *Microb. Pathogen.* 1992; 13: 261-9, and Wani, et al. *Infect. Immun.* 1996; 64:3967-3974). Antibodies to these adhesins could inhibit binding of pneumococci to respiratory epithelial cells and thereby reduce colonization. Other cytosolic pneumococcal proteins such as pneumolysin, autolysin, neuraminidase, or hyaluronidase are proposed as vaccine antigens because antibodies could potentially block the toxic effects of these proteins in patients infected with *S. pneumoniae.* However, these proteins are typically not located on the surface of *S. pneumoniae*, rather they are secreted or released from the bacterium as the cells lyse and die (Lee et al. *Vaccine* 1994; 12:875-8 and Berry et al. *Infect. Immun.* 1994; 62:1101-1108). While use of these cytosolic proteins as immunogens might ameliorate late consequences of *S. pneumoniae* infection, antibodies to these proteins would neither promote pneumococcal death nor prevent initial or subsequent pneumococcal colonization.

A prototypic surface protein that is being tested as a pneumococcal

vaccine is the pneumococcal surface protein A (PspA). PspA is a heterogeneous protein of about 70-140 kDa. The PspA structure includes an alpha helix at the amino terminus, followed by a proline-rich sequence, and terminates in a series of 11 choline-binding repeats at the carboxy-terminus. Although much information regarding its structure is available, PspA is not structurally conserved among a variety of pneumococcal serotypes, and its function is entirely unknown (Yother et al. *J. Bacteriol.* 1992;174:601-9 and Yother *J. Bacteriol.* 1994;176:2976-2985). Studies have confirmed the immunogenicity of PspA in animals (McDaniel et al. *Microb. Pathogen.* 1994; 17;323-337). Despite the immunogenicity of PspA, the heterogeneity of PspA, its existence in four structural groups (or clades), and its uncharacterized function complicate its ability to be used as a vaccine antigen.

In patients who cannot make protective antibodies to the type-specific polysaccharide capsule, the third component of complement, C3, and the associated proteins of the alternative complement pathway constitute the first line of host defense against *S. pneumoniae* infection. Because complement proteins cannot penetrate the rigid cell wall of *S. pneumoniae*, deposition of opsonic C3b on the pneumococcal surface is the principal mediator of pneumococcal clearance. Interactions of pneumococci with plasma C3 are known to occur during pneumococcal bacteremia, when the covalent binding of C3b, the opsonically active fragment of C3, initiates phagocytic recognition and ingestion (Johnston et al. *J. Exp. Med* 1969;129:1275-1290, Hasin HE, *J. Immunol.* 1972; 109:26-31 and Hostetter et al. *J. Infect. Dis.* 1984; 150:653-61). C3b deposits on the pneumococcal capsule, as well as on the cell wall. This method for controlling *S. pneumoniae* infection is fairly inefficient. Methods for augmenting *S. pneumoniae* opsonization could improve the disease course induced by this organism. There currently exists a strong need for methods and therapies to limit *S. pneumoniae* infection.

Summary of the Invention

This invention relates to the identification and use of a family of human complement C3-degrading proteins (proteinases) expressed by *S. pneumoniae*.

The proteins preferably have a molecular weight of about 15 kD to about 25 kD, as determined, for example, on a 10% SDS polyacrylamide gel. The invention includes a number of proteins isolatable from different C3-degrading strains of *S. pneumoniae*.

In one aspect, the invention relates to an isolated protein having at least 80% sequence identity with SEQ ID NO:2. In a preferred embodiment, the protein is isolated from *S. pneumoniae* or alternatively the protein is a recombinant protein. Preferably, the isolated protein degrades human complement protein C3. A preferred protein of this invention is an isolated protein having an amino acid sequence that includes SEQ ID NO:2, and more preferably, is SEQ ID NO:2. The term "isolated" as used herein refers to a naturally occurring species that has been removed from its natural environment, as well as to synthetic species. The term "protein" as used herein includes one or more functional units, which encompasses one or more peptides or polypeptides.

The invention also relates to isolated peptides or polypeptides from the C3-degrading proteinase of this invention. Preferably, the invention provides peptides or polypeptides of at least 15 sequential amino acids from an isolated protein that has at least 80% sequence identity with SEQ ID NO:2, and more preferably, peptides or polypeptides of at least 15 sequential amino acids of SEQ ID NO:2. In another aspect of this invention, the peptides or polypeptides are capable of degrading human complement protein C3.

Preferred embodiments of the invention include an isolated protein comprising amino acids of about 1 to about 58 of SEQ ID NO:2 and an isolated nucleic acid fragment comprising nucleotides of about 1 to about 174 of SEQ ID NO:1 or its complementary strand. Preferably, the isolated nucleic acid fragment comprises nucleotides of about 150 to about 174 of SEQ ID NO:1 or its complementary strand.

In another aspect, the invention relates to an isolated protein that degrades human complement protein C3, wherein nucleic acid encoding the protein hybridizes to SEQ ID NO:1 or its complementary strand under highly stringent hybridization conditions.

The invention also relates to an immune system stimulating composition (preferably, a vaccine) comprising an effective amount of an immune system stimulating peptide or polypeptide comprising at least 15 sequential amino acids derived from a protein, wherein the protein has at least 80% sequence identity with SEQ ID NO:2 and is capable of degrading human complement protein C3.

Preferably the protein is isolated from *S. pneumoniae*. In one embodiment, the immune system stimulating composition or vaccine further comprises at least one other immune system stimulating peptide, polypeptide or protein isolated from *S. pneumoniae*.

The invention further relates to an antibody capable of binding (typically, specifically binding) to a protein comprising at least 80% sequence identity with SEQ ID NO:2 and capable of degrading human complement protein C3. In one embodiment, the antibody is a monoclonal antibody and in another embodiment, the antibody is a polyclonal antibody. In another embodiment the antibody is an antibody fragment. The antibody or antibody fragments can be obtained from a mouse, a rat, a goat, a chicken, a human, or a rabbit.

In another embodiment, the antibody is capable of binding to at least a portion of a protein, wherein nucleic acid encoding the protein hybridizes to SEQ ID NO:1 or its complementary strand under highly stringent hybridization conditions.

The invention also relates to an isolated nucleic acid fragment (polynucleotide) capable of hybridizing to SEQ ID NO:1 or its complimentary strand under highly stringent hybridization conditions. As used herein, highly stringent hybridization conditions include, for example, 6XSSC, 5X Denhardt, 0.5% SDS, and 100 μg/ml fragmented and denatured salmon sperm DNA hybridized overnight at 65°C and washed in 2X SSC, 0.1% SDS one time at room temperature for about 10 minutes followed by one time at 65°C for about 15 minutes followed by at least one wash in 0.2XSSC, 0.1% SDS at room temperature for at least 3-5 minutes. In one embodiment, the nucleic acid fragment is isolated from *S. pneumoniae* and in another embodiment, the nucleic acid fragment encodes at least a portion of a protein. In one embodiment, the protein degrades human complement protein C3. In another embodiment, the

nucleic acid fragment encodes a peptide or polypeptide that does not degrade human complement C3.

In another embodiment, the nucleic acid fragment is in a nucleic acid vector and the vector can be an expression vector capable of producing at least a portion of a protein. Cells containing the nucleic acid fragment are also contemplated in this invention. In one embodiment, the cell is a bacterium or a eukaryotic cell.

The invention further relates to an isolated nucleic acid fragment comprising the nucleic acid sequence of SEQ ID NO:1, or its complementary strand. The invention further relates to an RNA fragment transcribed by a double-stranded DNA sequence comprising SEQ ID NO:1.

In another aspect of this invention, the invention relates to a method for producing an immune response to *S. pneumoniae* in a mammal (particularly a human) including the steps of: administering a composition comprising a therapeutically effective amount of at least a portion of a protein to an animal, wherein nucleic acid encoding the protein hybridizes to SEQ ID NO:1, or its complementary strand, under highly stringent hybridization conditions to produce an immune response to the protein. The immune response can be a B cell response, a T cell response, an epithelial response or an endothelial response. In a preferred embodiment, the composition is a vaccine composition. Preferably the protein is at least 15 amino acids in length and also preferably the composition further comprises at least one other immune system stimulating peptide, polypeptide or protein from *S. pneumoniae.* In one embodiment, the protein comprises at least 15 amino acids of SEQ ID NO:2.

The invention further relates to an isolated protein of about 15 kDa to about 25 kDa from *Streptococcus pneumoniae* that is capable of degrading human complement C3 and to a method for inhibiting *Streptococcus pneumoniae*-mediated C3 degradation comprising the step of: contacting a *Streptococcus pneumoniae* bacterium with antibody capable of binding to a protein with at least 80% amino acid sequence identity to SEQ ID NO:2.

The invention also relates to a method for inhibiting C3-mediated inflammation and rejection in xenotransplantation comprising the step of

expressing on the surface of an organ of an animal used in xenotransplantation a protein with the amino acid sequence of SEQ ID NO:2. This method is particularly advantageous for causing, for example, the kidneys of pigs to express the protein described herein and thereby to inhibit C3 mediated inflammation after xenotransplantation.

The invention also relates to an isolated nucleic acid molecule that contains a region of at least 15 nucleotides which hybridize under highly stringent hybridization conditions to at least a portion of a nucleic acid sequence of SEQ ID NO.1 or its complementary strand. In one embodiment, an isolated nucleic acid molecule is capable of hybridizing under highly stringent hybridization conditions to at least one region of SEQ ID NO.1 or its complementary strand. Preferably, the at least one region includes nucleotides 1-174 or 320-492 of SEQ ID NO:1.

In yet another embodiment, the invention relates to an isolated nucleic acid molecule that contains a region of at least 15 nucleotides which hybridize under highly stringent hybridization conditions to at least a portion of a nucleic acid sequence of SEQ ID NO.4 or its complementary strand. In one embodiment, an isolated nucleic acid molecule is capable of hybridizing under highly stringent conditions to at least one region of SEQ ID NO.4 or its complementary strand. Preferably, the at least one region includes nucleotides 507-681 or 827-999 of SEQ ID NO:4.

In another embodiment, at least a portion of the nucleic acid molecule of SEQ ID NO:4 encodes at least a portion of a protein. Preferably, the protein has a predicted amino acid sequence as shown in SEQ ID NO:5, and has a molecular weight of about 75 kDa to about 85 kDa as determined, for example, by SDS-PAGE.

The invention also relates to isolated DNA fragments or primers having the nucleic acid sequences as shown in SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8 AND SEQ ID NO:9.

In another aspect of the invention, the invention relates to an immune system-stimulating composition or vaccine containing a therapeutically effective amount of at least a portion of a protein, wherein nucleic acid encoding the

protein is capable of hybridizing to the sequence of SEQ ID NO:4, or its complementary strand, under highly stringent hybridization conditions. A method for producing an immune response to *S. pneumoniae* in a mammal (particularly a human) includes the steps of: administering a composition comprising a therapeutically effective amount of at least a portion of a protein to a mammal, wherein nucleic acid encoding the protein hybridizes to SEQ ID NO:4, or its complementary strand, under stringent hybridization conditions to obtain an immune response to the protein.

In yet another embodiment, the invention relates to a vaccine or immune system stimulating composition containing an amount of at least a portion of a protein that is effective to immunize or treat a mammalian subject against *S. pneumoniae* infection or colonization and a pharmaceutically acceptable carrier. The protein is derived from a nucleic acid molecule that hybridizes under highly stringent conditions to a nucleic acid sequence as shown in SEQ ID NO:4, or its complementary strand, coding for the protein. Preferably, the protein is provided in an amount effective to provide a therapeutic effect to the mammalian subject, especially a human subject.

Brief Description of the Figures

Figure 1 provides the nucleic acid sequence of the translated portion of a C3-degrading proteinase gene of this invention (SEQ ID NO:1).

Figure 2 provides the amino acid sequence of a C3-degrading proteinase of this invention (SEQ ID NO:2).

Figure 3 diagrams the amino acid sequence of a C3-degrading proteinase positioned with the nucleic acid sequence encoding a C3-degrading proteinase according to this invention (SEQ ID NOS:2-3).

Figure 4 provides the nucleic acid sequence for a predicted 79 kDa amino acid sequence (SEQ ID NO:4).

Figure 5 provides the predicted 79 kDa amino acid sequence (SEQ ID NO:5).

Figure 6 shows sequence alignments of SEQ ID NO:1 and SEQ ID NO:4.

Figure 7 shows sequence alignment of SEQ ID NO:2 with corresponding amino acids 169-331of SEQ ID NO:5.

Detailed Description of the Preferred Embodiments

The present invention relates to the identification and isolation of a human complement C3 degrading proteinase with a molecular weight of about 20 kDa (± 5 kDa) on a 10% SDS-PAGE gel and nucleic acid encoding a C3 degrading proteinase. It has been observed that exponentially growing cultures of pneumococci from several serotypes were able to first degrade the β-chain then degrade the α chain of C3 without producing defined C3 cleavage fragments (Angel, et al. *J. Infect. Dis.* 170:600-608, 1994). This pattern of degradation without cleavage differs substantially from that of other microbial products such as the elastase enzyme of *Pseudomonas aeruginosa* and the cysteine proteinase of *Entamoeba histolytica.*

The term "degrade" is used herein to refer to the removal of amino acids from proteinaceous molecules, generating peptides or polypeptides. The proteins of this invention degrade C3 without producing specific cleavage fragments as observed on a polyacrylamide gel. There is at least some preference of the C3-degrading proteinases of this invention for C3 in that, for example, the C3-degrading proteinase does not appear to degrade other proteins, such as albumin.

A C3-degrading proteinase of about 20 kDa was isolated from a library of insertionally interrupted pneumococcal genes by identifying those clones that had decreased C3 degrading activity as compared to wild type *S. pneumoniae*. An exemplary assay for assessing C3-degrading activity of clones is provided in Example 1. Clones with decreased C3-degrading activity were identified and a 546 bp SmaI insert was selected, based on the sequence of the clones that had demonstrated decreased C3-degrading activity. This SmaI fragment was used to probe a S. *pneumoniae* library made from strain CP1200. Positive clones from the *S. pneumoniae* library that hybridized to the SmaI fragment were isolated and the open reading frame of the gene associated with C3-degrading activity was identified. The following oligonucleotide (SEQ ID NO:10), which has sequence

identify with a portion of PspA, was used to confirm, by differential hybridization, that the gene encoding the C3-degrading protinease was distinct from the gene encoding PspA.

SEQ ID NO:10
GAAAACAATAATGTAGAAGACTACTTTAAAGAAGGTTAGA

A complete open reading frame of a 20 kDa protein spans an area of 492 base pairs (SEQ ID NO:1) predicting a protein of molecular weight of about 20 kDa (+/- 5 kDa) or about 163 amino acids (SEQ ID NO:2). An exemplary gene sequence encoding a C3-degrading protein is provided in Figure 1 as SEQ ID NO:1 and an amino acid sequence of the protein is provided in Figure 2 as SEQ ID NO:2. Figure 3 combines a preferred gene sequence with a corresponding preferred translated protein as SEQ ID NO:3.

Using SEQ ID NO:2, the amino acid sequence of the protein was determined to be unrelated to other proteins in the GenBank or Swiss Prot databases. The predicted protein encompasses a proline-rich sequence characteristic of membrane domains in prokaryotes, particularly between amino acids 80-108 suggesting that the protein is expressed at the surface. The amino acid sequence exhibits no apparent choline-binding repeats. Electrophoresis of pneumococcal lysates and supernatants from cultures of CP 1200 on SDS-PAGE gels impregnated with C3 identified a lytic band at about 20 kDa (± 5 kDa) in both supernatants and lysates, confirming that a protein of a size predicted by SEQ ID NO:2 had C3-degrading activity (see Example 2). As provided in Example 3, the gene encoding the 20 kDa C3-degrading proteinase is conserved in at least two dozen pneumococcal isolates representing five serotypes (serotypes 1, 3, 4, 14, and 19F).

The full length gene encoding a C3-degrading proteinase of this invention was inserted into a gene expression vector for expression in *E. coli*. Recombinant C3-degrading proteinase was isolated as described in the examples. Those of ordinary skill in the art recognize that, given a particular gene sequence such as that provided in SEQ ID NO:1, there are a variety of

expression vectors that could be used to express the gene. Further, there are a variety of methods known in the art that could be used to produce and isolate the recombinant protein of this invention and those of ordinary skill in the art also recognize that the C3 degrading assay of this invention will determine whether or not a particular expression system, in addition to those expression systems provided in the examples, is functioning, without requiring undue experimentation. A variety of molecular and immunological techniques can be found in basic technical texts such as those of Sambrook et al. (*Molecular Cloning, A Laboratory Manual*, 1989 Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY) and Harlow et al. (*Antibodies; A Laboratory Manual.* Cold Spring Harbor, NY; Cold Spring Harbor Laboratory Press, 1988).

The gene encoding the C3 degrading protein of this invention was identified using a plasmid library made with pneumococcal genomic DNA fragments from strain CP1200. Although there are a variety of methods known for obtaining a plasmid library, in a preferred strategy, a plasmid library was constructed with Sau 3A digested pneumococcal genomic DNA fragments (0.5 -4.0 kb) from pneumococcal strain CP 1200 (obtained from D.A. Morrison, University of Illinois, Champagne-Urbana, Illinois and described in Havarstein LF, et al. *Proc. Natl. Acad. Sci. (USA)* 1995;92:11140-11144) and inserted into the Bam HI site of the integrative shuttle vector pVA 891 (erm^r, cm^r; has origin of replication for *E. coli*). This library was transformed into an *E. coli* DH5α MCR strain by electroporation. Plasmid extractions of some randomly selected *E. coli* transformants revealed that all of them contained recombinant plasmids.

Plasmid library DNA can be extracted from the *E. coli* transformants and used to transform the CP 1200 parent pneumococcal strain using insertional mutagenesis by homologous recombination.

The pneumococcal strain CP 1200 cells can be made competent using a pH shift with HCl procedure in CTM medium. The competent cells are frozen at -70° C in small aliquots until needed.

The isolated protein of this invention can be incubated with human complement C3 for 4 hours at 37°C in the presence of PBS to detect C3 degradation. Control samples without the isolated pneumococcal protein are

used as controls for comparative purposes.

The protein of this invention has an apparent molecular weight on a 10% SDS-polyacrylamide gel of about 20 kDa (±5 kDa) and preferably has a molecular weight of about 15 kDa to about 25 kDa. An exemplary protein sequence is provided by SEQ ID NO: 2. Those of ordinary skill in the art will recognize that some variability in amino acid sequence is expected and that this variability should not detract from the scope of this invention. For example, conserved mutations do not detract from this invention nor do variations in amino acid sequence identity of less than about 80 % amino acid sequence identity and preferably less than about 90% amino acid sequence identity where the protein is capable of degrading human complement protein C3, and particularly where the protein is isolated or originally obtained from an *S. pneumoniae* bacterium. Fragments of the protein are also within the scope of the present invention, particularly if they are capable of degrading human complement protein C3.

Some nucleic acid sequence variability is expected among pneumococcal strains and serotypes as is some amino acid variability. Conserved amino acid substitutions are known in the art and include, for example, amino acid substitutions using other members from the same class to which the amino acid belongs. For example, the nonpolar amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, and tryptophan. The polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine and glutamine. The positively charged (basic) amino acids include arginine, lysine and histidine. The negatively charged (acidic) amino acids include aspartic acid and glutamic acid. Such alterations are not expected to affect apparent molecular weight as determined by polyacrylamide gel electrophoresis or isoelectric point. Particularly preferred conservative substitutions include, but are not limited to, Lys for Arg and *vice verse* to maintain a positive charge; Glu for Asp and vice versa to maintain a negative charge; Ser for Thr so that a free -OH is maintained; and Gln for Asn to maintain a free NH_2.

A preferred protein of this invention includes a protein with the amino acid sequence of SEQ ID NO:2. Other proteins include those degrading human

complement protein C3 and having nucleic acid encoding the protein that hybridizes to SEQ ID NO:1 under highly stringent hybridization conditions such as 6XSSC, 5X Denhardt, 0.5% SDS (sodium dodecyl sulfate), and 100 μg/ml fragmented and denatured salmon sperm DNA hybridized overnight at 65°C and washed in 2X SSC, 0.1% SDS one time at room temperature for about 10 minutes followed by one time at, 65°C for about 15 minutes followed by at least one wash in 0.2XSSC, 0.1% SDS at room temperature for at least 3-5 minutes are also contemplated in this invention. Typically, an SSC solution contains sodium chloride, sodium citrate, and water to prepare a stock solution. Peptides or polypeptides of the protein can also be used. A preferred protein of this invention comprises amino acids of about 1 to about 50 of SEQ ID NO:2.

The proteins of this invention can be isolated or prepared as recombinant proteins. That is, nucleic acid encoding the protein, or a portion of the protein, can be incorporated into an expression vector or incorporated into a chromosome of a cell to express the protein in the cell. The protein can be purified from a bacterium or another cell, preferably a eukaryotic cell and more preferably an animal cell. Alternatively, the protein can be isolated from a cell expressing the protein, such as a *S. pneumoniae* cell. Peptides or polypeptides are also considered in this invention. The peptides or polypeptides are preferably at least 15 amino acids in length and preferred peptides or polypeptides have at least 15 sequential amino acids from SEQ ID NO:2.

Nucleic acid encoding the 20 kDa protein is also part of this invention. SEQ ID NO:1 is a preferred nucleic acid fragment encoding a C3-degrading proteinase. Those of ordinary skill in the art will recognize that some substitution will not alter the C3-degrading proteinase sequence to an extent that the character or nature of the C3-degrading proteinase is substantially altered. For example, nucleic acid with an identity of at least 80% to SEQ ID NO:1 is contemplated in this invention. A method for determining whether a particular nucleic acid sequence falls within the scope of this invention is to consider whether or not a particular nucleic acid sequence encodes a C3-degrading proteinase and has a nucleic acid identity of at least 80% as compared with SEQ ID NO:1. Other nucleic acid sequences encoding the C3 proteinase include

nucleic acid encoding the C3 proteinase where the C3 proteinase has the same sequence or at least a 90% sequence identity with SEQ ID NO:2 but which includes degeneracy with respect to the nucleic acid sequence. A degenerate codon means that a different three letter codon is used to specify the same amino acid. For example, it is well known in the art that the following RNA codons (and therefore, the corresponding DNA codons, with a T substituted for a U) can be used interchangeably to code for each specific amino acid:

Phenylalanine (Phe or F)	UUU, UUC, UUA or UUG
Leucine (Leu or L)	CUU, CUC, CUA or CUG
Isoleucine (Ile or I)	AUU, AUC or AUA
Methionine (Met or M)	AUG
Valine (Val or V)	GUU, GUC, GUA, GUG
Serine (Ser or S)	AGU or AGC
Proline (Pro or P)	CCU, CCC, CCA, CCG
Threonine (Thr or T)	ACU, ACC, ACA, ACG
Alanine (Ala or A)	GCU, GCG, GCA, GCC
Tryptophan (Trp)	UGG
Tyrosine (Tyr or Y)	UAU or UAC
Histidine (His or H)	CAU or CAC
Glutamine (Gln or Q)	CAA or CAG
Asparagine (Asn or N)	AAU or AAC
Lysine (Lys or K)	AAA or AAG
Aspartic Acid (Asp or D)	GAU or GAC
Glutamic Acid (Glu or E)	GAA or GAG
Cysteine (Cys or C)	UGU or UGC
Arginine (Arg or R)	AGA or AGG
Glycine (Gly or G)	GGU or GGC or GGA or GGG
Termination codon	UAA, UAG or UGA

Further, a particular DNA sequence can be modified to employ the codons preferred for a particular cell type. For example, the preferred codon usage for *E. coli* is known, as are preferred codons for animals including humans. These changes are known to those of ordinary skill in the art and

therefore these gene sequences are considered part of this invention. Other nucleic acid sequences include nucleic acid fragments of at least 15, and preferably, at least 30 nucleic acids in length from SEQ ID NO:1 or other nucleic acid fragments of at least 15, and preferably at least 30 nucleic acids in length where these fragments hybridize to SEQ ID NO:1 under highly stringent hybridization conditions such as 6XSSC, 5X Denhardt, 0.5% SDS, and 100 μg/ml fragmented and denatured salmon sperm DNA hybridized overnight at 65°C and washed in 2X SSC, 0.1% SDS one time at room temperature for about 10 minutes followed by one time at, 65°C for about 15 minutes followed by at least one wash in 0.2XSSC, 0.1% SDS at room temperature for at least 3-5 minutes.

The nucleic acid fragments of this invention can encode all, none (i.e., fragments that cannot be transcribed, fragments that include regulatory portions of the gene, or the like) or a portion of SEQ ID NO:2 or SEQ ID NO:5 and preferably containing a contiguous nucleic acid fragment that encodes at least nine amino acids from SEQ ID NO:2 or SEQ ID NO:5. Because nucleic acid fragments encoding a portion of the C3 proteinase are contemplated in this invention, it will be understood that not all of the nucleic acid fragments will encode a protein or peptide or polypeptide with C3 degrading activity. Further, the nucleic acid of this invention can be mutated to remove or otherwise inactivate the C3 degrading activity of this protein. Therefore, fragments without C3 degrading activity that meet the hybridization requirements described above are also contemplated. Methods for mutating or otherwise altering nucleic acid sequences are well described in the art and the production of an immunogenic, but enzymatically inactive protein can be tested for therapeutic utility.

The nucleic acid fragments of this invention can be incorporated into nucleic acid vectors or stably incorporated into host genomes to produce recombinant protein including recombinant chimeric protein. In one embodiment, the C3-degrading protein is encoded by a gene in a vector and the vector is in a cell. Preferably, the cell is a prokaryotic cell such as a bacterium. The genes and gene fragments can exist as the fusion of all or a portion of the

gene with another gene and the C3-degrading protein can exist as a fusion protein of one or more proteins where the fusion protein is expressed as a single protein. A variety of nucleic acid vectors of this invention are known in the art and include a number of commercially available expression plasmids or viral vectors. The use of these vectors is well within the scope of what is ordinary skill in the art. Exemplary vectors are employed in the examples, but should not be construed as limiting on the scope of this invention.

This invention also relates to antibodies capable of binding (typically specifically binding) to a protein of about 20 kDa, and preferably a protein of about 15 kDa to about 25 kDa, from *S. pneumoniae* and preferably where the protein is capable of degrading human complement C3. Polyclonal antibody can be prepared to a portion of the protein or to all of the protein. Similarly, monoclonal antibodies can be prepared to all or to a peptide or polypeptide (fragment) of the about 20 kDa C3 degrading protein of this invention. Methods for preparing antibodies to protein are well known and well described, for example, by Harlow et al., (*Antibodies; A Laboratory Manual.* Cold Spring Harbor, NY; Cold Spring Harbor Laboratory Press, 1988). In a preferred example, the antibodies can be human derived, rat derived, mouse derived, goat derived, chicken derived, or rabbit derived. Protein-binding antibody fragments and chimeric fragments are also known and are within the scope of this invention.

The invention also relates to the use of immune stimulating compositions. The term "immune stimulating" or "immune system stimulating" composition refers to protein, peptide or polypeptide compositions according to the invention that activate at least one cell type of the immune system in a subject, such as a mammal. Preferably, the immune stimulating composition provides an immunizing response or prophylactic benefit in a normal, i.e., uninfected subject, typically a vaccine. However, any measurable immune response is beneficial to the subject in a therapy application or protocol. Preferred activated cells of the immune system include phagocytic cells such as neutrophils or macrophages, T cells, B cells, epithelial cells and endothelial cells. Immune stimulating compositions comprising the peptides, polypeptides

or proteins of the invention can be used to produce antibody in an animal such as a rat, mouse, goat, chicken, rabbit, or a human or an animal model for studying *S. pneumoniae* infection. Preferred immune stimulating compositions include an immune stimulating amount, e.g, a therapeutically effective amount, of at least one peptide or polypeptide including at least 15 amino acids from the C3 degrading proteinase.

The term "vaccine" refers to a composition for immunization. This process can include the administration of a protein, peptide, polypeptide, antigen, nucleic acid sequence or complementary sequence, e.g., anti-sense, or antibody, or suspensions thereof, wherein upon administration, the molecule will produce active immunity and provide protection against an *S. pneumoniae* infection or colonization. Typically, such vaccines are prepared as injectables, either as liquid solutions or suspensions. Solid forms suitable for solution in, or suspension in, liquid prior to injection may also be prepared. The vaccine preparation may optionally be emulsified, or encapsulated in liposomes.

The immune stimulating composition (such as a vaccine) can further include other proteins in a pharmaceutically acceptable buffer or carrier, such as PBS (phosphate buffer saline) or another buffer recognized in the art as suitable and safe for introduction of proteins into a host to stimulate the immune system. The immune stimulating compositions can also include other immune system stimulating proteins such as adjuvants or immune stimulating proteins, peptides or polypeptides from *S. pneumoniae* or other organisms. For example, a cocktail of peptides or polypeptides may be most useful for controlling *S. pneumoniae* infection. Preferably one or more fragments of the proteins of this invention are used in a vaccine preparation to protect against or limit *S. pneumoniae* colonization or the pathogenic consequences of *S. pneumoniae* colonization.

By a "therapeutically effective amount," as used herein, refers to that amount that is effective for production of a desired result. This amount varies depending upon the health and physical condition of a subject's immune system, i.e., to synthesize antibodies, the degree of protection desired, the formulation prepared and other relevant factors. It is expected that the amount will fall in a relatively broad range that can be determined through routine trials.

The active immune stimulating ingredients are often mixed with excipients or diluents that are pharmaceutically acceptable as carriers and compatible with the active ingredient. The term "pharmaceutically acceptable carrier" refers to a carrier(s) that is "acceptable" in the sense of being compatible with the other ingredients of a composition and not deleterious to the recipient thereof. Suitable excipients are, for example, water, saline, dextrose, glycerol, ethanol, or the like and combinations thereof. In addition, if desired, the immune stimulating composition (including vaccine) may contain minor amounts of auxiliary substances such as wetting or emulsifying agents, pH buffering agents, and/or adjuvants which enhance the effectiveness of the immune stimulating compostion.

Examples of adjuvants or carriers that may be effective include but are not limited to: aluminum hydroxide, N-acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP); N-acetyl-nor-muramyl-L-alanyl-D-isoglutamine (CGP 11637, referred to as nor-MDP), N-acetylmuramyl-L-alanyl-D-isoglutaminyl-L-alanine-2-(1'-2'-dipalmitoyl-sn-glycero-3-hydroxyphosphoryloxy)-ethylamine (CGP 19835A, referred to as MTP-PE), and RIBI, which contains three components extracted from bacteria, monophosphoryl lipid A, trehalose dimycolate and cell wall skeleton (MPL + TDM + CWS) in a 2% squalene/Tween 80 emulsion.

This invention also relates to a method for inhibiting *Streptococcus pneumoniae*-mediated C3 degradation comprising contacting a *Streptococcus pneumonie* bacterium with a protein, such as an antibody or another protein that is capable of binding to an isolated protein of about 15 kDa to about 25 kDa from *Streptococcus pneumoniae*. The protein capable of binding to an isolated protein of about 15 kDa to about 25 kDa can be an antibody or a fragment thereof or the protein can be a chimeric protein that includes the antibody binding domain, such as a variable domain, from antibody that is capable of specifically recognizing an isolated protein of about 15 kDa to about 25 kDa from *Streptococcus pneumoniae* having C3 degrading activity.

The isolated *S. pneumoniae* protein of this invention can be isolated, and optionally purified, and the isolated protein or immunogenic fragments thereof

can be used to produce an immunologic response, including, in one example, an antibody response in a human or an experimental animal. Peptide or polypeptide fragments of the protein without C3 degrading ability can be tested for their ability to limit the effects of *S. pneumoniae* infection. Similarly, the protein of this invention can be modified, such as through mutation to interrupt or inactivate the C3 degrading activity of the protein. Antibody capable of inhibiting the C3-degrading activity of the protein of this invention may be used as a strategy for preventing C3 degradation and for promoting clearance of *S. pneumoniae* through the opsonic pathway. Isolated protein can be used in assays to detect antibody to *S. pneumoniae* or as part of a vaccine or a multi-valent or multiple protein, peptide or polypeptide-containing vaccine for *S. pneumoniae* therapy.

Thus, the term "treatment," as used herein, refers to prophylaxis and/or therapy of either normal mammalian subjects or mammalian subjects colonized with, diagnosed with, or exhibiting characteristics or symptoms of various *S. pneumoniae* infections. The term "therapy" refers to providing a therapeutic effect to a mammalian subject such that the subject exhibits few or no symptoms of a pneumococcal infection or other related disease. Such treatment can be accomplished by administration of nucleic acid molecules (sense or antisense), proteins, peptides or polypeptides or antibodies of the instant invention.

It is further contemplated that the proteins of this invention can be surface expressed on vertebrate cells and used to degrade C3, for example, where complement deposition (or activation) becomes a problem, such as in xenotransplantation or in complement-mediated glomerulonephritis. For example, the entire pneumococcal protein, a recombinant protein, or a portion of either, can be incorporated into xenotransplant cells and expressed as a surface protein or as a secreted protein to prevent or minimize complement deposition (and/or complement-mediated inflammation).

Another specific aspect of the present invention relates to using a vaccine vector expressing an isolated protein and peptides or polypeptides therefrom. Accordingly, in a further aspect this invention provides a method of inducing an immune response in a mammal, which comprises providing to a mammal a

vaccine vector expressing at least one, or a mixture of an isolated protein and/or peptide or polypeptide of the invention. The protein and peptides or polypeptides of the present invention can be delivered to the mammal using a live vaccine vector, in particular using live recombinant bacteria, viruses or other live agents, containing the genetic material necessary for the expression of the protein and/or peptides or polypeptides as a foreign polypeptide. Particularly, bacteria that colonizes the gastrointestinal tract, such as Salmonella, Shigella, Yersinia, Vibrio, Escherichia and BCG have been developed as vaccine vectors, and these and other examples are discussed by J. Holmgren et al., Immunobiol., 184, 157-179 (1992) and J. McGhee et al., Vaccine, 10, 75-88 (1992).

An additional embodiment of the present invention relates to a method of inducing an immune response in a subject, e.g,. mammal, comprising administering to the subject an amount of a DNA molecule encoding an isolated protein and/or peptide or polypeptide therefrom of this invention, optionally with a transfection-facilitating agent, where the protein and/or peptides or polypeptides retain immunogenicity and, when incorporated into an immune stimulating composition, e.g, vaccine, and administered to a human, provides protection without inducing enhanced disease upon subsequent infection of the human with *S. pneumoniae* pathogen. Transfection-facilitating agents are known in the art.

It is further contemplated that the antisense sequence of the gene encoding the 20kDa protein may be used as a vaccine or as a therapeutic treatment for pneumococcal infection. Antisense DNA is defined as a non-coding sequence that is complementary, i.e., a complementary strand, to all or a portion of SEQ ID NO:1. For example, the antisense sequence for 5'-ATGTCAAGC-3' is 3'-TACAGTTCG-5'. Delivery of antisense sequence or oligonuclotides into an animal may result in the production of antibody by the animal or in the incorporation of the sequence into living bacteria or other cells whereby transcription and/or translation of all or a portion of the 20 kDa gene product is inhibited.

Introduction of an antisense nucleic acid sequence can be accomplished, for example, by loading the antisense nucleic acid into a suitable carrier, such as

a liposome, for introduction into pneumococci or infected cells. Typically, an antisense nucleic acid sequence having eight or more nucleotides is capable of binding to the bacterial nucleic acid or bacterial messenger RNA. The antisense nucleic acid sequence, typically contains at least about 15 nucleotides, preferably at least about 30 nucleotides or more nucleotides to provide necessary stability of a hybridization product of bacterial nucleic acid or bacterial messenger RNA. Introduction of the sequences preferably inhibit the transcription or translation of at least one endongenous *S. pneumoniae* nucleic acid sequence. Methods for loading antisense nucleic acid is known in the art as exemplified by U.S. Patent 4,242,046.

The present invention also provides nucleic acid having an open reading frame of 2163 bases (SEQ ID NO:4) that encompasses the open reading frame of a nucleic acid sequence (SEQ ID NO:1) that encodes a protein that has a molecular weight of about 20 kDa (SEQ ID NO:2). The 20 kDa protein, described herein, is further characterized as a C3-degrading protein. The larger open reading frame, e.g., 2163 bp (SEQ ID NO:4), encodes for a putative protein of about 79 kDa (SEQ ID NO:5).

All references and publications cited herein are expressly incorporated by reference into this disclosure. There are a variety of alternative techniques and procedures available to those of skill in the art which would similarly permit one to successfully perform the intended invention in view of the present disclosure. It will be appreciated by those skilled in the art that while the invention has been described above in connection with particular embodiments and examples, the invention is not necessarily so limited and that numerous other embodiments, examples, uses, modifications and departures from the embodiments, examples and uses may be made without departing from the inventive scope of this application.

Example 1

Identification of Insertional Mutants with Reduced C3-Degrading Activity

Insertional mutants were received from Dr. Elaine Tuomanen, (Rockefeller Inst., New York, New York). The clones with insertions were tested in an assay to detect reduced C3-degrading activity. 137 clones were tested by growing the cells in Todd Hewitt broth overnight at room temperature in microtitre plates. The cells were diluted 1:10 in synthetic medium for pneumococci (see Sicard A. M., *Genetics* 50:31-44, 1984) and the remainder of the cells were frozen in the microtiter plate. Either 63 ng or 83 ng of C3 per 100 µl of medium containing 1 mg/ml of 0.1% BSA in phosphate buffered saline (PBS) was added to about 200 µl of diluted cells. The cells were incubated at 37°C for 4 hrs. One hundred µl of the mixture was added to ELISA plates and incubated overnight at 4°C. The plates were washed three times with wash buffer and the wells were filled with 0.05% Tween 20 in PBS with five minute incubations between the washes. One hundred µl of antibody to C3 (polyclonal horse-radish peroxidase-conjugated goat antibody specific to human C3-IgG fraction, ICN Cappel, Costa Mesa, CA) was diluted 1:1200 with 3% BSA in PBS. The ELISA plate was incubated at 37°C for about 30 minutes to 1 hr in the dark and washed with wash buffer as above. The assay was developed using 12 mg of OPD in 30 ml of 0.1M sodium citrate buffer with 12 µl of 30% hydrogen peroxide. Assay results were determined by optical density readings at 490 nm on an ELISA plate reader.

Each clone was tested four times. Nineteen clones were selected that had less than 40% C3 degradation as compared to nonmutated controls. These 19 clones were screened 6 times by the assay described above and from these results 6 clones were selected with less than 30% C3-degrading activity as compared to controls. These 6 clones were screened eleven times each and the two clones with the lowest C3-degrading activity were selected for further study.

A partial sequence of one of the clones was received and a *SmaI* fragment of 546bp was labeled with ^{32}P by random primer labeling (kit available

from Stratagene, La Jolla, CA). The 546bp SmaI fragment from SEQ ID NO:1 was hybridized to EcoR1 and KpnI digests of numerous pneumococcal strains on Southern blots. This same fragment was also used to screen a library of Sau3A fragments of genomic DNA from S. pneumoniae strain CP1200.

A 3.5 kb insert was identified from the CP1200 library. The insert was sequenced and an open reading frame of 492 base pairs, including the stop codon, was identified. The open reading frame coded for a protein of 163 amino acids and a predicted molecular weight of about 18,500 daltons.

PCR primers were constructed to amplify the open reading frame; the 5' PCR primer incorporated a *BamHI* site; the 3' primer incorporated a *PstI* site. The amplified insert was ligated in frame to a His-Tagged *E. coli* expression vector pQE30 (Qiagen, San Diego, CA). The resulting plasmid was used to transform *E. coli* strain BL21 (Novagen, Madison, WI) containing the lac repressor plasmid pREP4 (Qiagen). *E. coli* cultures were induced to express the His-Tagged protein and the protein was column purified with Ni-NTA resin (Qiagen). The purified protein was confirmed by SDS-PAGE gel.

Example 2

Identification of a 20kDa C3-Degrading Proteinase

To determine the C3-degrading capability of the 20 kDa protein, 0.5 mg/ml of C3 (prepared according to Tack et al., *Meth. Enzymol.* 80:64-101, 1984) was copolymerized in a sodium dodecyl sulfate (SDS) gel-containing 15% acrylamide (15% SDS-PAGE gel). Pneumococcal supernatants were obtained from cultures of S. pneumoniae strain CP1200 grown to exponential phase in Todd Hewitt broth; pneumococcal lysates were obtained by incubating 5×10^8 cells with 5% SDS for 30 minutes at room temperature. The lysate was concentrated 10 fold using a Centricon filtration device with a 10,000 mw cutoff (Amicon, Beverly, MA). The samples were not heated before electrophoresis. Samples of supernatants and lysates were added to the 15% C3-containing SDS-PAGE gels and electrophoresis was carried out at 4°C at 150 V until the dye front ran out. The gel was washed successively with 50 ml of 2.5% Triton X-100 in water (2 times, 10 minutes), 2.5 % Triton X-100 in 50 mM Tris-HCl, pH

7.4 (2 times, 10 minutes), and 50 mm Tris-HCl, pH 7.4 (2 times, 10 minutes) to remove SDS. After washes, 50 ml of 50 mM Tris-HCl, pH 7.4, was poured into dishes containing the gels, and the dishes were covered and incubated at 37°C for 1.5 hour and overnight (about 16 hours). The gels were stained with Coomassie blue for 10 minutes and destained totally.

Two lytic bands were visualized, one of which was about 20 kDa in size, against the dark blue background in both lysates and supernatant. C3 proteinase activities in the pneumococcal lysates were observed after a 1.5 hour incubation at 37°C, while C3 proteinase activities in the Pn supernatant were observed after an overnight incubation. Therefore, C3 proteinase activities appeared to be mainly cell associated.

Example 3

The gene encoding the 20 kD protein is conserved in a number of *S. pneumoniae* strains.

DNA was obtained from a variety of *S. pneumoniae* strains (Clinical isolates of Type 1, Type 3, LOO2 and LOO3 (type 3), Type 4, Type 14 and Laboratory isolates CP1200, WU2, R6X, 6303,109,110, JY1119, JY182, and JY53) and SEQ ID NO:3 was used as a probe to detect the presence of nucleic acid encoding the 20 kD protein in DNA from these strains. Isolated chromosomal DNA was digested with *EcoRI* and separated by electrophoresis. The DNA was transferred to a solid support and hybridized to end-labeled SEQ ID NO:3 under the hybridization and washing conditions of 6X SSC, 5X Denhart's, 0.5% SDS, 100 µg/ml denatured fragmented salmon sperm DNA hybridized at 65 °C overnight and washed in 2X SSC, 1 time at room temperature for 10 minutes and in 2X SSC, 0.1% SDS 1 time at 65°C for 15 minutes followed by two washes in 0.2X SSC, 0.1% SDS for 3 minutes each at room temperature.

Results indicated that SEQ ID NO:3 hybridized identically in each of the DNA samples tested indicating that the protein appears to be conserved among strains. In some strains, the DNA encoding the 20 kDa C3-degrading protein

appears to be part of a larger open reading frame of 2163 bp that putatively encodes a 79 kDa protein.

Example 4

Southern blot of S. pneumoniae DNA/5F1 probe

Five ug samples of genomic DNA were obtained from 11 strains *of S. pneumoniae*. Each sample was digested with the restriction enzyme Kpn1. The samples were subsequently loaded onto an agarose gel and resolved by electrophoresis. The samples contained in the gel were subsequently transferred to a Hybond-N+ membrane available from Amersham (Upsalla, Sweden) by capillary transfer. A 540bp SmaI fragment from an 5F1 isolate was random prime labeled with P^{32} using a ^{T7}QuickPrime kit (Pharmacia, Piscathaway, NJ) and purified from non-incorporated nucleotides using NucTrap column (Stragene, La Jolla, CA) and hybridized.

The hybridization conditions were 6XSSC, 5X Denhardt, 0.5% SDS, and 100 µg/ml fragmented and denatured salmon sperm DNA hybridized overnight at 65°C and washed in 2X SSC, 0.1% SDS once at room temperature for about 10 minutes followed by 1 time at, 65°C for about 15 minutes followed by at least one wash in 0.2XSSC, 0.1% SDS at room temperature for at least 3-5 minutes. The blot demonstrated that the 20kDa gene was present in all tested strains of *S. pneumoniae*.

Example 5

Two DNA primers were prepared from SEQ ID NO:1 and utilized to amplify the 20Kda gene sequence from *S. pneumoniae* (serotype 3) genomic DNA. The first primer, a 5'-primer, SEQ ID NO:6, spans the ATG start codon of the 20Kda gene, inserts a Nco1 site, and had an Ala residue inserted after the ATG start codon to maintain a correct reading frame. The second primer, a 3'-primer, SEQ ID NO:7, spans the termination codon of the 20 Kda gene and inserts a BamH1 site.

5'-GGGGG <u>CCA TGG</u> CC TCA AGC CTT TTA CGT GAA TTG-3'; (SEQ ID NO:6)

5'-GGGGG <u>GGA TCC</u> CTA GCT ATA TGA GAT AAA CTT TCC TGC T-3'; (SEQ ID NO:7)

The two primers were synthesized on an Applied Biosystems 380A DNA synthesizer (Foster City, CA) using reagents purchased from Glen Research (Sterling, VA). Amplifications were performed utilizing a Perkin Elmer Thermocycler (ABI) according to the manufacturer's directions. The identified PCR product was ligated into the TA tailed PCR cloning vector PCR2.1, available from Invitrogen, Carlsbad, CA, and used to transform OneShot Top10F' competent cells (Invitrogen). Kanamycin resistant transformants were screened by restriction enzyme analysis of plasmid DNA prepared by alkaline lysis. An approximately 500bp insert fragment was identified and subsequently excised with restriction enzymes Nco1 and BamH1. The 500 bp fragment was purified from a low melting agarose gel, and subsequently ligated into the Nco1-BamH1 sites of the T7 promoted expression vector pET 28a, available from Novagen (Madison, WI).

The ligation mixture was subsequently transformed into Top10F'cells (Invitrogen), and the kanamycin resistant transformants were screened by restriction enzyme analysis of plasmid DNA prepared by alkaline lysis. A recombinant plasmid (pLP505) was subsequently tranformed into BL21 (Novagen) cells and grown in SOB media supplemented with 30ug/ml kanamycin. Cells were grown to an O.D.$_{600}$ of 0.6, and were subsequently

induced with 0.4mM IPTG (Boehringer Mannheim, Indianapolis, Indiana) for 2-4 hours. Whole cell lysates were prepared and electrophoresed on a 14% SDS-PAGE gel. The gel was stained with Coomassie and the expression product was detected. The coomassie stained gel revealed a band between the 28 kDa and the 18 kDa molecular weight markers, and was determined to be approximately 20 kDa.

The DNA sequence of the insert in the recombinant pLP505 plasmid was obtained using the ABI 370A DNA sequencer. The DNA sequence was aligned with the DNA sequence of SEQ ID NO:1, using the Pustell DNA matrix plot feature of MacVector (Oxford Molecular Group, Campbell, CA). Alignment of the DNA sequence obtained from the pLP505 plasmid, SEQ ID NO:1, and the *S. pneumoniae* (serotype 4) genome, revealed that the open reading frame (ORF) that codes for the 20Kd protein may be part of a larger ORF, i.e., a 2163 bp in the serotype 4 genome, that codes for a protein with a predicted MW of approximately 79 kDa (SEQ ID NO: 4). DNA SEQ ID NO:4 encodes for a predicted amino acid sequence as shown in SEQ ID NO:5.

The *S. pneumoniae* (serotype 4) genome sequence was obtained from The Institute for Genomic Research at www.tigr.org and/or through NCBI at www.ncbi.nlm.nih.gov, using the ClustalW feature of MacVector, (Oxford Molecular Group, Campbell CA). A sequence comparison was made between the 20 kDa amino acid sequence (SEQ ID NO:2) and the predicted 79 kDa amino acid sequence (SEQ ID NO:5). It was observed that amino acids 1-58 and amino acids 90-132 of SEQ ID NO:2 have substantial sequence identity with amino acids 170-227 and amino acids 258-300 of SEQ ID NO:5, respectively. Proteins and peptides or polypeptides, containing these particular regions are preferred embodiments of the invention.

Based upon the available genomic DNA (serotype 4) sequence, two primers flanking the 2163bp ORF were designed and subsequently synthesized using the ABI 380A DNA synthesizer (SEQ ID NOS:8 and 9). SEQ ID NO:8 was an *S. pneumoniae* 5'-primer having an inserted Nco1 site and a "Glu" residue added after the ATG start codon to maintain a correct reading frame. SEQ ID NO:9, was an *S. pneumoniae* 3'-primer having an inserted HindIII site.

5'-AGA GCT CCT CCC ATG GAA GAT CCG AAT TAT CAG-3'; (SEQ ID NO:8)

5'-CCG GGC AAG CTT TTA CTT ACT CTC CT-3'-; (SEQ ID NO:9)

An approximately 2100 bp DNA fragment was then amplified from the 4 different S. pneumoniae serotypes (serotype 3, 5, 6B and 7) resulting in 4 fragments. Each of the 4 fragments were subsequently ligated into the PCR cloning vector PCR2.1 (Invitrogen), and used to transform OneShot Top 10F' cells (Invitrogen). Kanamycin resistant transformants were screened by restriction analysis of the plasmid DNA prepared by alkaline lysis. A recombinant plasmid containing the serotype 7 PCR product was identified, e.g., pLP512. The DNA sequence was obtained from the serotype 7 clone using the ABI model 370A DNA sequencer. The DNA sequence was essentially identical to SEQ ID NO:4 and encoded a predicted amino acid sequence essentially identical to SEQ ID NO:5.

It will be appreciated by those skilled in the art that while the invention has been described above in connection with particular embodiments and examples, the invention is not necessarily so limited and that numerous other embodiments, examples, uses, modifications and departures from the embodiments, examples and uses may be made without departing from the inventive scope of this application.

What is claimed is:

1. An isolated protein having at least 80% sequence identity with SEQ ID NO:2 which is capable of degrading human complement protein C3.

2. The protein of claim 1 which is isolated from *S. pneumoniae*.

3. The protein of claim 1 which is a recombinant protein.

4. The protein of claim 1 having a molecular weight of about 15 kDa to about 25 kDa.

5. An isolated peptide or polypeptide comprising at least 15 sequential amino acids from the protein of claim 1.

6. An isolated peptide or polypeptide comprising at least 15 sequential amino acids of SEQ ID NO:2.

7. An isolated protein comprising SEQ ID NO:2.

8. The isolated protein of claim 7 which has a molecular weight of about 15 kDa to about 25 kDa.

9. The protein of claim 8 which is isolated from *S. pneumoniae*.

10. The protein of claim 8 which degrades human complement protein C3.

11. The protein of claim 7 which is SEQ ID NO:2.

12. An isolated protein comprising amino acids from about 1 to about 58 of SEQ ID NO:2.

13. The isolated protein of claim 12 further comprising amino acids from about 90 to about 132 of SEQ ID NO:2.

14. An isolated protein comprising amino acids from about 170 to about 227 of SEQ ID NO:5.

15. The isolated protein of claim 14 further comprising amino acids from about 258 to about 300 of SEQ ID NO:5.

16. An isolated protein that degrades human complement protein C3, wherein nucleic acid encoding the protein hybridizes to SEQ ID NO:1 or its complementary strand under highly stringent hybridization conditions.

17. An isolated protein of about 15 kDa to about 25 kDa from *S. pneumoniae* that is capable of degrading human complement C3.

18. An immune system stimulating composition comprising an effective amount of an immune system stimulating peptide or polypeptide comprising at least 15 sequential amino acids derived from a protein, wherein the protein has at least 80% sequence identity with SEQ ID NO:2 and is capable of degrading human complement protein C3.

19. The immune system stimulating composition of claim 18 wherein the protein is isolated from *S. pneumoniae.*

20. The immune system stimulating composition of claim 19 further comprising at least one other immune system stimulating protein, peptide or polypeptide isolated from *S. pneumoniae*.

21. An immune system stimulating composition comprising a therapeutically effective amount of at least a portion of a protein, wherein nucleic acid

encoding the protein hybridizes to SEQ ID NO:1 or its complementary strand under highly stringent hybridization conditions.

22. An immune system stimulating composition comprising an effective amount of at least a portion of the protein of claim 17 effective to immunize or treat a mammalian subject against *S. pneumoniae* infection or colonization, and a pharmaceutically acceptable carrier.

23. The immune system stimulating composition of claim 22 wherein the protein is provided in an amount effective to provide a therapeutic effect to the mammalian subject.

24. An antibody capable of binding to a protein comprising at least 80% sequence identity with SEQ ID NO:2 and capable of degrading human complement protein C3.

25. The antibody of claim 24 which is a monoclonal antibody.

26. The antibody of claim 24 which is obtained from a mouse, a rat, a goat, a chicken, a human, or a rabbit.

27. An antibody capable of binding to at least a portion of a protein, wherein nucleic acid encoding the protein hybridizes to SEQ ID NO:1 or its complementary strand under highly stringent hybridization conditions.

28. An isolated nucleic acid fragment capable of hybridizing to SEQ ID NO:1 or its complementary strand under highly stringent hybridization conditions.

29. The nucleic acid fragment of claim 28 isolated from *S. pneumoniae*.

30. The nucleic acid fragment of claim 28 wherein the nucleic acid fragment encodes at least a portion of a protein.

31. The nucleic acid fragment of claim 30 wherein the protein degrades human complement C3.

32. The nucleic acid fragment of claim 28 in a nucleic acid vector.

33. The nucleic acid fragment of claim 32 wherein the vector is an expression vector capable of producing at least a portion of a protein.

34. A cell comprising the nucleic acid of claim 28.

35. The cell of claim 34 wherein the cell is a bacterium or a eukaryotic cell.

36. An isolated nucleic acid fragment comprising nucleotides of about 1 to about 174 of SEQ ID NO:1 or its complementary strand.

37. The isolated nucleic acid fragment of claim 36 further comprising nucleotides of about 320 to about 492 of SEQ ID NO:1 or its complementary strand.

38. An isolated nucleic acid fragment comprising the nucleic acid sequence of SEQ ID NO:1 or its complementary strand.

39. An RNA fragment transcribed by a double-stranded DNA sequence comprising SEQ ID NO:1 or its complementary strand.

40. A method for producing an immune response to *S. pneumoniae* in a mammal comprising the steps of:

administering a composition comprising a therapeutically effective amount of at least a portion of a protein to a mammal,

wherein nucleic acid encoding the protein hybridizes to SEQ ID NO:1 or its complementary strand under highly stringent hybridization conditions in a pharmaceutically acceptable carrier to yield an immune response.

41. The method of claim 40 wherein the immune response is a B cell response, a T cell response, an epithelial cell response, or an endothelial cell response.

42. The method of claim 40 wherein the at least a portion of a protein is at least 15 amino acids in length.

43. The method of claim 40 wherein the composition further comprises at least one other immune system stimulating protein, peptide or polypeptide from *S. pneumoniae*.

44. The method of claim 40 wherein the at least a portion of a protein comprises at least 15 amino acids of SEQ ID NO:2.

45. A method for inhibiting *S. pneumoniae*-mediated C3 degradation comprising the step of:

contacting a *S. pneumoniae* bacterium with an antibody capable of binding to a protein with the amino acid sequence of SEQ ID NO:2 or a fragment thereof.

46. A method for inhibiting C3-mediated inflammation and rejection in xenotransplantation comprising the step of:

expressing on the surface of an organ of an animal used in xenotransplantation a protein with the amino acid sequence of SEQ ID NO:2 or a fragment thereof.

47. An isolated nucleic acid molecule comprising a region of at least 15

nucleotides which hybridize under highly stringent hybridization conditions to at least a portion of a nucleic acid sequence as shown in SEQ ID NO:1, or its complementary strand.

48. An isolated nucleic acid molecule comprising a sequence that hybridizes under highly stringent hybridization conditions to at least one region of SEQ ID NO:1 or its complementary strand wherein the region is selected from the group consisting of nucleotides 1-174 and 320-492.

49. An isolated nucleic acid molecule comprising a region of at least 15 nucleotides which hybridize under highly stringent hybridization conditions to at least a portion of a nucleic acid sequence as shown in SEQ ID NO:4, or its complementary strand.

50. An isolated nucleic acid molecule comprising a sequence that hybridizes under highly stringent hybridization conditions to at least one region of SEQ ID NO:4 or its complementary strand, wherein the region is selected from the group consisting of nucleotides 507-681 and 827-999.

51. The nucleic acid molecule of claim 49 which encodes at least a portion of a protein.

52. The nucleic acid molecule of claim 51 wherein the protein has a predicted amino acid sequence of SEQ ID NO:5.

53. An isolated nucleic acid fragment having a nucleic acid sequence selected from the group consisting of SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, AND SEQ ID NO:9.

54. An immune system stimulating composition comprising a therapeutically effective amount of at least a portion of a protein, wherein nucleic acid encoding the protein hybridizes to SEQ ID NO:4 or its complementary

strand under highly stringent hybridization conditions.

55. A method for producing an immune response to *S. pneumoniae* in a mammal comprising the steps of:

administering a composition comprising a therapeutically effective amount of at least a portion of a protein to a mammal, wherein nucleic acid encoding the protein hybridizes to SEQ ID NO:4, or its complementary strand under highly stringent hybridization conditions in a pharmaceutically acceptable carrier to yield an immune response.

56. An immune system stimulating composition comprising an effective amount of at least a portion of the protein of claim 51 effective to immunize or treat a mammalian subject against *S. pneumoniae* infection or colonization, and a pharmaceutically acceptable carrier.

57. The immune system stimulating composition of claim 56 wherein the protein is provided in an amount effective to provide a therapeutic effect to the mammalian subject.

58. The immune system stimulating composition of claim 56 wherein the composition is a vaccine.

59. A polypeptide having SEQ ID NO:5.

60. The immune system stimulating composition of claim 23 where the protein encoded by the nucleic acid sequence or its complementary strand inhibits the transcription or translation of at least one endogenous *S. pneumoniae* nucleic acid sequence.

61. The immune system stimulating composition of claim 56 where the protein encoded by the nucleic acid sequence or its complementary

strand inhibits the transcription or translation of at least one endogenous *S. pneumoniae* nucleic acid sequence.

1 ATGTCAAGCC TTTTACGTGA ATTGTATGCT AAACCCTTAT CAGAACGCCA
51 TGTAGAATCT GATGGTCTTA TTTTCGACCC AGCGCAAATC ACAAGTCGAA
101 CCGCCAATGG TGTTGCTGTA CCGCACGGAG ACCATTATCA CTTTATTCCT
151 TATTCACAAC TGTCACCTTT GGAAGAAAAA TTGGTCGTAT TATTCCCCTT
201 CGTTATCGTT CAAACCATTG GGTACCAGAT TCAAAGACCA GAACAACCAG
251 TCCACAATCG ACTCCGGGAA CCTAGTCCAA GTCCGAAACC TGCACCAAAT
301 CCTCAACCAG CTCCAAGCAA TCCAATTGAT GAGAAATTGG TCAAAGAAGC
351 TGTTCGAAAA GTAGGCGATG GTTATGTCTT TGAGGAGAAT GGAGTTGCCT
401 CGTTATATCC CAAGCCAAGG ATCTTACAGC AGAAACAGCA GCAGGCATTG
451 ATAGCAAACT GGCCAAGCAG GAAAGTTTAT CTCATAAGCT AG

Fig. 1

(SEQ ID NO: 1)

1 MSSLLRELYA KPLSERHVES DGLIFDPAQI TSRTANGVAV PHGDHYHFIP
51 YSQLSPLEEK LVVLFPFVIV QTIGYQIQRP EQPVHNRLRE PSPSPKPAPN
101 PQPAPSNPID EKLVKEAVRK VGDGYVFEEN GVASLYPKPRILQQKQQQAL
151 LANWPSRKVY LIS*

Fig. 2

(SEQ ID NO: 2)

```
    ATGTCAAGCCTTTTACGTGAATTGTATGCTAAACCCTTATCAGAACGCCATGTAGAATCT
  1 ---------+---------+---------+---------+---------+---------+ 60
    TACAGTTCGGAAAATGCACTTAACATACGATTTGGGAATAGTCTTGCGGTACATCTTAGA

    M  S  S  L  L  R  E  L  Y  A  K  P  L  S  E  R  H  V  E  S  -

    GATGGTCTTATTTTCGACCCAGCGCAAATCACAAGTCGAACCGCCAATGGTGTTGCTGTA
 61 ---------+---------+---------+---------+---------+---------+ 120
    CTACCAGAATAAAAGCTGGGTCGCGTTTAGTGTTCAGCTTGGCGGTTACCACAACGACAT

    D  G  L  I  F  D  P  A  Q  I  T  S  R  T  A  N  G  V  A  V  -

    CCGCACGGAGACCATTATCACTTTATTCCTTATTCACAACTGTCACCTTTGGAAGAAAAA
121 ---------+---------+---------+---------+---------+---------+ 180
    GGCGTGCCTCTGGTAATAGTGAAATAAGGAATAAGTGTTGACAGTGGAAACCTTCTTTTT

    P  H  G  D  H  Y  H  F  I  P  Y  S  Q  L  S  P  L  E  E  K  -

    TTGGTCGTATTATTCCCCTTCGTTATCGTTCAAACCATTGGGTACCAGATTCAAAGACCA
181 ---------+---------+---------+---------+---------+---------+ 240
    AACCAGCATAATAAGGGGAAGCAATAGCAAGTTTGGTAACCCATGGTCTAAGTTTCTGGT

    L  V  V  L  F  P  F  V  I  V  Q  T  I  G  Y  Q  I  Q  R  P  -

    GAACAACCAGTCCACAATCGACTCCGGGAACCTAGTCCAAGTCCGAAACCTGCACCAAAT
241 ---------+---------+---------+---------+---------+---------+ 300
    CTTGTTGGTCAGGTGTTAGCTGAGGCCCTTGGATCAGGTTCAGGCTTTGGACGTGGTTTA

    E  Q  P  V  H  N  R  L  R  E  P  S  P  S  P  K  P  A  P  N  -
```

Fig. 3

```
    CCTCAACCAGCTCCAAGCAATCCAATTGATGAGAAATTGGTCAAAGAAGCTGTTCGAAAA
301 ---------+---------+---------+---------+---------+---------+ 360
    GGAGTTGGTCGAGGTTCGTTAGGTTAACTACTCTTTAACCAGTTTCTTCGACAAGCTTTT

    P  Q  P  A  P  S  N  P  I  D  E  K  L  V  K  E  A  V  R  K  -

    GTAGGCGATGGTTATGTCTTTGAGGAGAATGGAGTTGCCTCGTTATATCCCAAGCCAAGG
361 ---------+---------+---------+---------+---------+---------+ 420
    CATCCGCTACCAATACAGAAACTCCTCTTACCTCAACGGAGCAATATAGGGTTCGGTTCC

    V  G  D  G  Y  V  F  E  E  N  G  V  A  S  L  Y  P  K  P  R  -

    ATCTTACAGCAGAAACAGCAGCAGGCATTGATAGCAAACTGGCCAAGCAGGAAAGTTTAT
421 ---------+---------+---------+---------+---------+---------+ 480
    TAGAATGTCGTCTTTGTCGTCGTCCGTAACTATCGTTTGACCGGTTCGTCCTTTCAAATA

    I  L  Q  Q  K  Q  Q  Q  A  L  I  A  N  W  P  S  R  K  V  V  -

    CTCATAAGCTAG
481 ---------+-- 492
    GAGTATTCGATC

    L  I  S  *  -
```

Fig 3 (CONTINUED)

ATGAAAGATCCGAATTATCAGTTGAAGGATTCAGACATTGTCAAT
GAAATCAAGGGTGGTTATGTTATCAAGGTAGATGGAAAATACTA
TGTTTACCTTAAGGATGCAGCTCATGCGGATAATATTCGGACAAA
AGAAGAGATTAAACGTCAGAAGCAGGAACACAGTCATAATCACG
GGGGTGGTTCTAACGATCAAGCAGTAGTTGCAGCCAGAGCCCAA
GGACGCTATACAACGGATGATGGTTATATCTTCAATGCATCTGAT
ATCATTGAGGACACGGGTGATGCTTATATCGTTCCTCACGGCGAC
CATTACCATTACATTCCTAAGAATGAGTTATCAGCTAGCGAGTTA
GCTGCTGCAGAAGCCTATTGGAATGGGAAGCAGGGATCTCGTCCT
TCTTCAAGTTCTAGTTATAATGCAAATCCAGCTCAACCAAGATTG
TCAGAGAACCACAATCTGACTGTCACTCCAACTTATCATCAAAAT
CAAGGGGAAAACATTTCAAGCCTTTTACGTGAATTGTATGCTAAA
CCCTTATCAGAACGCCATGTGGAATCTGATGGCCTTATTTTCGAC
CCAGCGCAAATCACAAGTCGAACCGCCAGAGGTGTAGCTGTCCC
TCATGGTAACCATTACCACTTTATCCCTTATGAACAAATGTCTGA
ATTGGAAAAACGAATTGCTCGTATTATTCCCCTTCGTTATCGTTCA
AACCATTGGGTACCAGATTCAAGACCAGAACAACCAAGTCCACA
ATCGACTCCGGAACCTAGTCCAAGTCCGCAACCTGCACCAAATCC
TCAACCAGCTCCAAGCAATCCAATTGATGAGAAATTGGTCAAAG
AAGCTGTTCGAAAAGTAGGCGATGGTTATGTCTTTGAGGAGAATG
GAGTTTCTCGTTATATCCCAGCCAAGGATCTTTCAGCAGAAACAG
CAGCAGGCATTGATAGCAAACTGGCCAAGCAGGAAAGTTTATCT
CATAAGCTAGGAGCTAAGAAAACTGACCTCCCATCTAGTGATCG
AGAATTTTACAATAAGGCTTATGACTTACTAGCAAGAATTCACCA
AGATTTACTTGATAATAAAGGTCGACAAGTTGATTTTGAGGCTTT
GGATAACCTGTTGGAACGACTCAAGGATGTCCCAAGTGATAAAG
TCAAGTTAGTGGATGATATTCTTGCCTTCTTAGCTCCGATTCGTCA
TCCAGAACGTTTAGGAAAACCAAATGCGCAAATTACCTACACTG
ATGATGAGATTCAAGTAGCCAAGTTGGCAGGCAAGTACACAACA
GAAGACGGTTATATCTTTGATCCTCGTGATATAACCAGTGATGAG
GGGGATGCCTATGTAACTCCACATATGACCCATAGCCACTGGATT
AAAAAAGATAGTTTGTCTGAAGCTGAGAGAGCGGCAGCCCAGGC
TTATGCTAAAGAGAAAGGTTTGACCCCTCCTTCGACAGACCATCA
GGATTCAGGAAATACTGAGGCAAAAGGAGCAGAAGCTATCTACA
ACCGCGTGAAAGCAGCTAAGAAGGTGCCACTTGATCGTATGCCTT
ACAATCTTCAATATACTGTAGAAGTCAAAAACGGTAGTTTAATCA

Fig. 4

TACCTCATTATGACCATTACCATAACATCAAATTTGAGTGGTTTG
ACGAAGGCCTTTATGAGGCACCTAAGGGGTATACTCTTGAGGATC
TTTTGGCGACTGTCAAGTACTATGTCGAACATCCAAACGAACGTC
CGCATTCAGATAATGGTTTTGGTAACGCTAGCGACCATGTTCAAA
GAAACAAAAATGGTCAAGCTGATACCAATCAAACGGAAAAACCA
AGCGAGGAGAAACCTCAGACAGAAAAACCTGAGGAAGAAACCC
CTCGAGAAGAGAAACCGCAAAGCGAGAAACCAGAGTCTCCAAAA
CCAACAGAGGAACCAGAAGAATCACCAGAGGAATCAGAAGAAC
CTCAGGTCGAGACTGAAAAGGTTGAAGAAAAACTGAGAGAGGCT
GAAGATTTACTTGGAAAAATCCAGGATCCAATTATCAAGTCCAAT
GCCAAAGAGACTCTCACAGGATTAAAAAATAATTTACTATTTGGC
ACCCAGGACAACAATACTATTATGGCAGAAGCTGAAAAACTATT
GGCTTTATTAAAGGAGAGTAAG

Fig. 4 (CONTINUED)

MKDPNYQLKDSDIVNEIKGGYVIKVDGKYYVYLKDAAHADNIRTK
EEIKRQKQEHSHNHGGGSNDQAVVAARAQGRYTTDDGYIFNASDII
EDTGDAYIVPHGDHYHYIPKNELSASELAAAEAYWNGKQGSRPSSS
SSYNANPAQPRLSENHNLTVTPTYHQNQGENISSLLRELYAKPLSER
HVESDGLIFDPAQITSRTARGVAVPHGNHYHFIPYEQMSELEKRIARII
PLRYRSNHWVPDSRPEQPSPQSTPEPSPSPQPAPNPQPAPSNPIDEKLV
KEAVRKVGDGYVFEENGVSRYIPAKDLSAETAAGIDSKLAKQESLS
HKLGAKKTDLPSSDREFYNKAYDLLARIHQDLLDNKGRQVDFEALD
NLLERLKDVPSDKVKLVDDILAFLAPIRHPERLGKPNAQITYTDDEIQ
VAKLAGKYTTEDGYIFDPRDITSDEGDAYVTPHMTHSHWIKKDSLSE
AERAAAQAYAKEKGLTPPSTDHQDSGNTEAKGAEAIYNRVKAAKK
VPLDRMPYNLQYTVEVKNGSLIIPHYDHYHNIKFEWFDEGLYEAPK
GYTLEDLLATVKYYVEHPNERPHSDNGFGNASDHVQRNKNGQADT
NQTEKPSEEKPQTEKPEEETPREEKPQSEKPESPKPTEEPEESPEESEEP
QVETEKVEEKLREAEDLLGKIQDPIIKSNAKETLTGLKNNLLFGTQD
NNTIMAEAEKLLALLKESK

Fig. 5

ClustalW Formatted Alignments

```
                      10        20        30
20Kseq      ATGTCAAGCCTTTTACGTGAATTGTATGCT
79KfragDNA  ATTTCAAGCCTTTTACGTGAATTGTATGCT

                      40        50        60
20Kseq      AAACCCTTATCAGAACGCCATGTAGAATCT
79KfragDNA  AAACCCTTATCAGAACGCCATGTGGAATCT

                      70        80        90
20Kseq      GATGGTCTTATTTTCGACCCAGCGCAAATC
79KfragDNA  GATGGCCTTATTTTCGACCCAGCGCAAATC

                     100       110       120
20Kseq      ACAAGTCGAACCGCCAATGGTGTTGCTGTA
79KfragDNA  ACAAGTCGAACCGCCAGAGGTGTAGCTGTC

                     130       140       150
20Kseq      CCGCACGGAGACCATTATCACTTTATTCCT
79KfragDNA  CCTCATGGTAACCATTACCACTTTATCCCT

                     160       170       180
20Kseq      TATTCACAACTGTCACCTTTGGAAGAA-AA
79KfragDNA  TATGAACAAATGTCTGAATTGGAAAAACGA

                     190       200       210
20Kseq      ATTGGTCGTATTATTCCCCTTCGTTATCGT
79KfragDNA  ATTGCTCGTATTATTCCCCTTCGTTATCGT
```

Fig. 6

(SEQ ID NOS: 1 and 4)

```
                                 220                 230               240
20Kseq       TCAAACCATTGGGTACCAGATTCAAAGACC
79KfragDNA   TCAAACCATTGGGTACCAGATTCAA-GACC

                                 250                 260               270
20Kseq       AGAACAACCA-GTCCACAATCGACTCCGGG
79KfragDNA   AGAACAACCAAGTCCACAATCGACTCCGG-

                                 280                 290               300
20Kseq       AACCTAGTCCAAGTCCGAAACCTGCACCAA
79KfragDNA   AACCTAGTCCAAGTCCGCAACCTGCACCAA

                                 310                 320               330
20Kseq       ATCCTCAACCAGCTCCAAGCAATCCAATTG
79KfragDNA   ATCCTCAACCAGCTCCAAGCAATCCAATTG

                                 340                 350               360
20Kseq       ATGAGAAATTGGTCAAAGAAGCTGTTCGAA
79KfragDNA   ATGAGAAATTGGTCAAAGAAGCTGTTCGAA

                                 370                 380               390
20Kseq       AAGTAGGCGATGGTTATGTCTTTGAGGAGA
79KfragDNA   AAGTAGGCGATGGTTATGTCTTTGAGGAGA

                                 400                 410               420
20Kseq       ATGGAGTTGCCTCGTTATATCCCAAGCCAA
79KfragDNA   ATGGAGTT-TCTCGTTATATCCCA-GCCAA

                                 430                 440               450
20Kseq       GGATCTTACAGCAGAAACAGCAGCAGGCAT
79KfragDNA   GGATCTTTCAGCAGAAACAGCAGCAGGCAT

                                 460                 470               480
20Kseq       TGATAGCAAACTGGCCAAGCAGGAAAGTTT
79KfragDNA   TGATAGCAAACTGGCCAAGCAGGAAAGTTT

                                 490                 500               510
20Kseq       ATCTCATAAGCTAG
79KfragDNA   ATCTCATAAGCTAG
```

Fig. 6 (CONTINUED)
(SEQ ID NOS: 1 and 4)

ClustalW Formatted Alignments

```
                     10        20        30
20Kseq       MSSLLRELYAKPLSERHVESDGLIFDPAQI
79KfragDNA   ISSLLRELYAKPLSERHVESDGLIFDPAQI

                     40        50        60
20Kseq       TSRTANGVAVPHGDHYHFIPYSQLSPLEEK
79KfragDNA   TSRTARGVAVPHGNHYHFIPYEQMSELEKR

                     70        80        90
20Kseq       LVVLFPFVIVQTIGYQIQRPEQPVHNRLRE
79KfragDNA   IARIIPLRYRSNHWVPDSRPEQPSPQSTPE

                    100       110       120
20Kseq       PSPSPKPAPNPQPAPSNPIDEKLVKEAVRK
79KfragDNA   PSPSPQPAPNPQPAPSNPIDEKLVKEAVRK

                    130       140       150
20Kseq       VGDGYVFEENGVASLYPKPRILQQKQQ--Q
79KfragDNA   VGDGYVFEENGVSRYIPAKDLSAETAAGID

                    160       170       180
20Kseq       ALLANWPSRKVYLIS
79KfragDNA   SKLAKQESLSHKL
```

Fig. 7

(SEQ ID NO:2 and a portion of SEQ ID NO: 5)

SUBSTITUTE SEQUENCE LISTING

<110> HOSTETTER, Margaret K.
FINKEL, David J.
CHENG, Qi
GREEN, Bruce A.
MASI, Amy W.
REGENTS OF THE UNIVERSITY OF MINNESOTA

<120> HUMAN COMPLEMENT C3-DEGRADING PROTEINASE FROM STREPTOCOCCUS PNEUMONIAE

<130> 11000570201 Human Complement

<140> Not Assigned

<141> 1998-09-24

<150> 60/059,907

<151> 1997-09-24

<160> 10

<170> PatentIn Ver. 2.0

<210> 1

<211> 492

<212> DNA

<213> Streptococcus pneumoniae

<400> 1

atgtcaagcc ttttacgtga attgtatgct aaacccttat cagaacgcca tgtagaatct 60

gatggtctta ttttcgaccc agcgcaaatc acaagtcgaa ccgccaatgg tgttgctgta 120

ccgcacggag accattatca ctttattcct tattcacaac tgtcaccttt ggaagaaaaa 180

ttggtcgtat tattcccctt cgttatcgtt caaaccattg ggtaccagat tcaaagacca 240

gaacaaccag tccacaatcg actccgggaa cctagtccaa gtccgaaacc tgcaccaaat 300

cctcaaccag ctccaagcaa tccaattgat gagaaattgg tcaaagaagc tgttcgaaaa 360

gtaggcgatg gttatgtctt tgaggagaat ggagttgcct cgttatatcc caagccaagg 420

atcttacagc agaaacagca gcaggcattg atagcaaact ggccaagcag gaaagtttat 480

ctcataagct ag 492

<210> 2

<211> 163

<212> PRT

<213> Streptococcus pneumoniae

<400> 2

```
Met Ser Ser Leu Leu Arg Glu Leu Tyr Ala Lys Pro Leu Ser Glu Arg
  1               5                  10                  15

His Val Glu Ser Asp Gly Leu Ile Phe Asp Pro Ala Gln Ile Thr Ser
             20                  25                  30

Arg Thr Ala Asn Gly Val Ala Val Pro His Gly Asp His Tyr His Phe
         35                  40                  45

Ile Pro Tyr Ser Gln Leu Ser Pro Leu Glu Glu Lys Leu Val Val Leu
     50                  55                  60

Phe Pro Phe Val Ile Val Gln Thr Ile Gly Tyr Gln Ile Gln Arg Pro
65                  70                  75                  80

Glu Gln Pro Val His Asn Arg Leu Arg Glu Pro Ser Pro Ser Pro Lys
                 85                  90                  95

Pro Ala Pro Asn Pro Gln Pro Ala Pro Ser Asn Pro Ile Asp Glu Lys
            100                 105                 110

Leu Val Lys Glu Ala Val Arg Lys Val Gly Asp Gly Tyr Val Phe Glu
        115                 120                 125

Glu Asn Gly Val Ala Ser Leu Tyr Pro Lys Pro Arg Ile Leu Gln Gln
    130                 135                 140

Lys Gln Gln Gln Ala Leu Ile Ala Asn Trp Pro Ser Arg Lys Val Tyr
145                 150                 155                 160

Leu Ile Ser
```

<210> 3

<211> 492

<212> DNA

<213> Streptococcus pneumoniae

<400> 3

ctagcttatg agataaactt tcctgcttgg ccagtttgct atcaatgcct gctgctgttt 60

ctgctgtaag atccttggct tgggatataa cgaggcaact ccattctcct caaagacata 120

accatcgcct acttttcgaa cagcttcttt gaccaatttc tcatcaattg gattgcttgg 180

agctggttga ggatttggtg caggtttcgg acttggacta ggttcccgga gtcgattgtg 240

gactggttgt tctggtcttt gaatctggta cccaatggtt tgaacgataa cgaaggggaa 300

taatacgacc aatttttctt ccaaaggtga cagttgtgaa taaggaataa agtgataatg 360

gtctccgtgc ggtacagcaa caccattggc ggttcgactt gtgatttgcg ctgggtcgaa 420

aataagacca tcagattcta catggcgttc tgataagggt ttagcataca attcacgtaa 480

aaggcttgac at 492

<210> 4

<211> 2163

<212> DNA

<213> Streptococcus pneumoniae

<400> 4

atgaaagatc cgaattatca gttgaaggat tcagacattg tcaatgaaat caagggtggt 60

tatgttatca aggtagatgg aaaatactat gtttacctta aggatgcagc tcatgcggat 120

aatattcgga caaaagaaga gattaaacgt cagaagcagg aacacagtca taatcacggg 180

ggtggttcta acgatcaagc agtagttgca gccagagccc aaggacgcta tacaacggat 240

gatggttata tcttcaatgc atctgatatc attgaggaca cgggtgatgc ttatatcgtt 300

cctcacggcg accattacca ttacattcct aagaatgagt tatcagctag cgagttagct 360

gctgcagaag cctattggaa tgggaagcag ggatctcgtc cttcttcaag ttctagttat 420

aatgcaaatc cagctcaacc aagattgtca gagaaccaca atctgactgt cactccaact 480

tatcatcaaa atcaagggga aaacatttca agccttttac gtgaattgta tgctaaaccc 540

ttatcagaac gccatgtgga atctgatggc cttattttcg acccagcgca aatcacaagt 600

cgaaccgcca gaggtgtagc tgtccctcat ggtaaccatt accactttat cccttatgaa 660

caaatgtctg aattggaaaa acgaattgct cgtattattc cccttcgtta tcgttcaaac 720

cattgggtac cagattcaag accagaacaa ccaagtccac aatcgactcc ggaacctagt 780

ccaagtccgc aacctgcacc aaatcctcaa ccagctccaa gcaatccaat tgatgagaaa 840

ttggtcaaag aagctgttcg aaaagtaggc gatggttatg tctttgagga gaatggagtt 900

tctcgttata tcccagccaa ggatctttca gcagaaacag cagcaggcat tgatagcaaa 960

ctggccaagc aggaaagttt atctcataag ctaggagcta agaaaactga cctcccatct 1020

agtgatcgag aattttacaa taaggcttat gacttactag caagaattca ccaagattta 1080

cttgataata aaggtcgaca agttgatttt gaggctttgg ataacctgtt ggaacgactc 1140

aaggatgtcc caagtgataa agtcaagtta gtggatgata ttcttgcctt cttagctccg 1200

attcgtcatc cagaacgttt aggaaaacca aatgcgcaaa ttacctacac tgatgatgag 1260

attcaagtag ccaagttggc aggcaagtac acaacagaag acggttatat ctttgatcct 1320

cgtgatataa ccagtgatga gggggatgcc tatgtaactc cacatatgac ccatagccac 1380

tggattaaaa aagatagttt gtctgaagct gagagagcgg cagcccaggc ttatgctaaa 1440

gagaaaggtt tgacccctcc ttcgacagac catcaggatt caggaaatac tgaggcaaaa 1500

ggagcagaag ctatctacaa ccgcgtgaaa gcagctaaga aggtgccact tgatcgtatg 1560

ccttacaatc ttcaatatac tgtagaagtc aaaaacggta gtttaatcat acctcattat 1620

gaccattacc ataacatcaa atttgagtgg tttgacgaag gcctttatga ggcacctaag 1680

gggtatactc ttgaggatct tttggcgact gtcaagtact atgtcgaaca tccaaacgaa 1740

cgtccgcatt cagataatgg ttttggtaac gctagcgacc atgttcaaag aaacaaaaat 1800

ggtcaagctg ataccaatca aacggaaaaa ccaagcgagg agaaacctca gacagaaaaa 1860

cctgaggaag aaacccctcg agaagagaaa ccgcaaagcg agaaaccaga gtctccaaaa 1920

ccaacagagg aaccagaaga atcaccagag gaatcagaag aacctcaggt cgagactgaa 1980

aaggttgaag aaaaactgag agaggctgaa gatttacttg gaaaaatcca ggatccaatt 2040

atcaagtcca atgccaaaga gactctcaca ggattaaaaa ataatttact atttggcacc 2100

caggacaaca atactattat ggcagaagct gaaaaactat tggctttatt aaaggagagt 2160

aag 2163

<210> 5

<211> 721

<212> PRT

<213> Streptococcus pneumoniae

<400> 5

Met Lys Asp Pro Asn Tyr Gln Leu Lys Asp Ser Asp Ile Val Asn Glu
1 5 10 15

Ile Lys Gly Gly Tyr Val Ile Lys Val Asp Gly Lys Tyr Tyr Val Tyr
20 25 30

Leu Lys Asp Ala Ala His Ala Asp Asn Ile Arg Thr Lys Glu Glu Ile
35 40 45

```
Lys Arg Gln Lys Gln Glu His Ser His Asn His Gly Gly Gly Ser Asn
     50                  55                  60

Asp Gln Ala Val Val Ala Ala Arg Ala Gln Gly Arg Tyr Thr Thr Asp
65                  70                  75                  80

Asp Gly Tyr Ile Phe Asn Ala Ser Asp Ile Ile Glu Asp Thr Gly Asp
                85                  90                  95

Ala Tyr Ile Val Pro His Gly Asp His Tyr His Tyr Ile Pro Lys Asn
            100                 105                 110

Glu Leu Ser Ala Ser Glu Leu Ala Ala Ala Glu Ala Tyr Trp Asn Gly
        115                 120                 125

Lys Gln Gly Ser Arg Pro Ser Ser Ser Ser Ser Tyr Asn Ala Asn Pro
    130                 135                 140

Ala Gln Pro Arg Leu Ser Glu Asn His Asn Leu Thr Val Thr Pro Thr
145                 150                 155                 160

Tyr His Gln Asn Gln Gly Glu Asn Ile Ser Ser Leu Leu Arg Glu Leu
                165                 170                 175

Tyr Ala Lys Pro Leu Ser Glu Arg His Val Glu Ser Asp Gly Leu Ile
            180                 185                 190

Phe Asp Pro Ala Gln Ile Thr Ser Arg Thr Ala Arg Gly Val Ala Val
        195                 200                 205

Pro His Gly Asn His Tyr His Phe Ile Pro Tyr Glu Gln Met Ser Glu
    210                 215                 220

Leu Glu Lys Arg Ile Ala Arg Ile Ile Pro Leu Arg Tyr Arg Ser Asn
225                 230                 235                 240

His Trp Val Pro Asp Ser Arg Pro Glu Gln Pro Ser Pro Gln Ser Thr
```

```
                245                 250                 255

Pro Glu Pro Ser Pro Ser Pro Gln Pro Ala Pro Asn Pro Gln Pro Ala
            260                 265                 270

Pro Ser Asn Pro Ile Asp Glu Lys Leu Val Lys Glu Ala Val Arg Lys
        275                 280                 285

Val Gly Asp Gly Tyr Val Phe Glu Glu Asn Gly Val Ser Arg Tyr Ile
    290                 295                 300

Pro Ala Lys Asp Leu Ser Ala Glu Thr Ala Ala Gly Ile Asp Ser Lys
305                 310                 315                 320

Leu Ala Lys Gln Glu Ser Leu Ser His Lys Leu Gly Ala Lys Lys Thr
                    325                 330                 335

Asp Leu Pro Ser Ser Asp Arg Glu Phe Tyr Asn Lys Ala Tyr Asp Leu
            340                 345                 350

Leu Ala Arg Ile His Gln Asp Leu Leu Asp Asn Lys Gly Arg Gln Val
        355                 360                 365

Asp Phe Glu Ala Leu Asp Asn Leu Leu Glu Arg Leu Lys Asp Val Pro
    370                 375                 380

Ser Asp Lys Val Lys Leu Val Asp Asp Ile Leu Ala Phe Leu Ala Pro
385                 390                 395                 400

Ile Arg His Pro Glu Arg Leu Gly Lys Pro Asn Ala Gln Ile Thr Tyr
                405                 410                 415

Thr Asp Asp Glu Ile Gln Val Ala Lys Leu Ala Gly Lys Tyr Thr Thr
            420                 425                 430

Glu Asp Gly Tyr Ile Phe Asp Pro Arg Asp Ile Thr Ser Asp Glu Gly
        435                 440                 445
```

Asp Ala Tyr Val Thr Pro His Met Thr His Ser His Trp Ile Lys Lys
450 455 460

Asp Ser Leu Ser Glu Ala Glu Arg Ala Ala Ala Gln Ala Tyr Ala Lys
465 470 475 480

Glu Lys Gly Leu Thr Pro Pro Ser Thr Asp His Gln Asp Ser Gly Asn
485 490 495

Thr Glu Ala Lys Gly Ala Glu Ala Ile Tyr Asn Arg Val Lys Ala Ala
500 505 510

Lys Lys Val Pro Leu Asp Arg Met Pro Tyr Asn Leu Gln Tyr Thr Val
515 520 525

Glu Val Lys Asn Gly Ser Leu Ile Ile Pro His Tyr Asp His Tyr His
530 535 540

Asn Ile Lys Phe Glu Trp Phe Asp Glu Gly Leu Tyr Glu Ala Pro Lys
545 550 555 560

Gly Tyr Thr Leu Glu Asp Leu Leu Ala Thr Val Lys Tyr Tyr Val Glu
565 570 575

His Pro Asn Glu Arg Pro His Ser Asp Asn Gly Phe Gly Asn Ala Ser
580 585 590

Asp His Val Gln Arg Asn Lys Asn Gly Gln Ala Asp Thr Asn Gln Thr
595 600 605

Glu Lys Pro Ser Glu Glu Lys Pro Gln Thr Glu Lys Pro Glu Glu Glu
610 615 620

Thr Pro Arg Glu Glu Lys Pro Gln Ser Glu Lys Pro Glu Ser Pro Lys
625 630 635 640

Pro Thr Glu Glu Pro Glu Glu Ser Pro Glu Glu Ser Glu Glu Pro Gln
645 650 655

```
Val Glu Thr Glu Lys Val Glu Glu Lys Leu Arg Glu Ala Glu Asp Leu
            660                 665                 670

Leu Gly Lys Ile Gln Asp Pro Ile Ile Lys Ser Asn Ala Lys Glu Thr
        675                 680                 685

Leu Thr Gly Leu Lys Asn Asn Leu Leu Phe Gly Thr Gln Asp Asn Asn
    690                 695                 700

Thr Ile Met Ala Glu Ala Glu Lys Leu Leu Ala Leu Leu Lys Glu Ser
705                 710                 715                 720

Lys
```

<210> 6

<211> 34

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<220>

<223> Incorporates a Nco1 site and a DNA codon for Ala.

<400> 6

gggggccatg gcctcaagcc ttttacgtga attg 34

<210> 7

<211> 39

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<220>

<223> Incorporates a BamH1 site.

<400> 7

ggggggatc cctagctata tgagataaac tttcctgct 39

<210> 8

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<220>

<223> Incorporates a Nco1 site and a DNA codon for Glu.

<400> 8

agagctcctc ccatggaaga tccgaattat cag 33

<210> 9

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<220>

<223> Incorporates a HinD III site.

<400> 9

ccgggcaagc ttttacttac tctcct 26

<210> 10

<211> 40

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: probe

<220>

<223> Oligonucleotide

<400> 10

gaaaacaata atgtagaaga ctactttaaa gaaggttaga 40

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 98/20186

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/57 C12N9/52 A61K38/48 C07K16/40 C12N1/21
C12N5/10

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	ANGEL ET AL: "Degradation of C3 by Streptococcus pneumoniae" JOURNAL OF INFECTIOUS DISEASES, vol. 170, no. 3, 1994, pages 600-608, XP002092872 cited in the application see the whole document ---	1-61
O,A	NANDIWADA ET AL: "Genetic Analysis of a C3 Degrading Proteinase in Streptococcus pneumoniae" 96TH GENERAL MEETING OF THE AMERICAN SOCIETY FOR MICROBIOLOGY, NEW ORLEANS, LOUISIANA, SA, MAY 19-23, 1996. ABSTRACTS OF THE GENERAL MEETING OF THE AMERICAN SOCIETY FOR MICROBIOLOGY 96 (0),1996, page 177 XP002076135 see abstract B-134 ---	1-61

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
12 February 1999	25/02/1999

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Jansen, K-S

3

International Application No
PCT/US 98/20186

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
P,X	WO 98 18930 A (HUMAN GENOME SCIENCES INC ;CHOI GIL H (US); HROMOCKYJ ALEX (US); J) 7 May 1998 see table 1, particularly SEQ ID:65, SEQ ID:66, SEQ ID:94 see abstract; claims ---	5,6,14, 15,18, 19,21, 28-30, 32-35, 40-42, 44,48-50
E	WO 98 48022 A (NANDIWADA LAKSHMI S ;DUNNY GARY (US); UNIV MINNESOTA (US); HOSTETT) 29 October 1998 see abstract; claims -----	17

3

International application No.

PCT/US 98/20186

INTERNATIONAL SEARCH REPORT

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
Remark: Although claims 45-46 (as far as in vivo methods are concerned), and claims 40-44, 55 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

..iformation on patent family members

Inte onal Application No
PCT/US 98/20186

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9818930 A	07-05-1998	AU 5194598 A	22-05-1998
		AU 6909098 A	22-05-1998
		WO 9818931 A	07-05-1998
WO 9848022 A	29-10-1998	AU 7156698 A	13-11-1998

Form PCT/ISA/210 (patent family annex) (July 1992)